APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/542,409	07/14/2005	Yumiko Kato	MTS-3576US	1733

52473   7590   08/18/2009

RATNERPRESTIA
P.O. BOX 980
VALLEY FORGE, PA 19482

EXAMINER
GUERRA-ERAZO, EDGAR X

ART UNIT	PAPER NUMBER
2626	

MAIL DATE	DELIVERY MODE
08/18/2009	PAPER

Please find below and/or attached an Office communication concerning this application or proceeding.

The time period for reply, if any, is set in the attached communication.

	Application No.	Applicant(s)	
Office Action Summary	10/542,409	KATO ET AL.	
	Examiner	Art Unit	
	EDGAR GUERRA-ERAZO	2626	

-- *The MAILING DATE of this communication appears on the cover sheet with the correspondence address* --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133).
  Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1) ☒ Responsive to communication(s) filed on <u>14 July 2005</u>.
2a) ☐ This action is FINAL.   2b) ☒ This action is non-final.
3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4) ☒ Claim(s) <u>1-37</u> is/are pending in the application.
     4a) Of the above claim(s) <u>1-12,15,18,21-24 and 35</u> is/are withdrawn from consideration.
5) ☐ Claim(s) _____ is/are allowed.
6) ☒ Claim(s) <u>13,14,16,17,19,20,25-34,36 and 37</u> is/are rejected.
7) ☐ Claim(s) _____ is/are objected to.
8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9) ☐ The specification is objected to by the Examiner.
10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
     Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
     Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
     a) ☐ All   b) ☐ Some * c) ☐ None of:
       1. ☐ Certified copies of the priority documents have been received.
       2. ☐ Certified copies of the priority documents have been received in Application No. _____.
       3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
     * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☒ Information Disclosure Statement(s) (PTO/SB/08)
     Paper No(s)/Mail Date <u>07/14/2005, 05/01/2006, 11/13/2006 and</u>
     <u>07/17/2008.</u>

4) ☐ Interview Summary (PTO-413)
     Paper No(s)/Mail Date. _____
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____

DETAILED ACTION

Introduction

1.   This office action is in response to Applicant's submission filed on 07/14/2005. Claims

are pending in the application and have been examined.

2.   In response to submission of the preliminary amendment filed on 07/14/2005, where in

the drawings, in Figures 29, 31, and 34-36, the word "means" has been changed to "part", the

Examiner acknowledges Applicant's amendment.

3.   In response to the submission of the preliminary amendment filed on   02/10/2006,

cancelling Claims 1-11, and adding Claims 12-37, the Examiner acknowledges Applicant's

amendment.

4.   In response to the submission of the further preliminary amendment filed on

10/16/2006,   amending Claim 32, and cancelling Claims 1-12, 15, 18, 21-24, and 35, the

Examiner acknowledges Applicant's amendment.

Information Disclosure Statement

5.   The Information Disclosure Statements filed on 07/14/2005, 05/01/2006, 11/13/2006 and

07/17/2008 have been accepted and considered in this office action.

Priority

6.   Acknowledgment is made of applicant's claim for foreign priority under 35 U.S.C.

119(a)-(d). The certified copy has been filed in the instant application.

Claim Rejections - 35 USC § 103

7.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in
> section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are
> such that the subject matter as a whole would have been obvious at the time the invention was made to a person
> having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negatived by the
> manner in which the invention was made.

8.   Claims 13, 14, 16, 17, 19, 20, 25-34, 36, 37 are rejected under 35 U.S.C. 103(a) as being

unpatentable over Nakatsuyama (U.S. Patent Application: 2002/0143550) in view of Mitchell et

al.,(U.S. Patent: 6,961,700); hereinafter referred to as Nakatsuyama and Mitchell.

9.   With respect to **Claims 13, 16, 19, 25, 36**, Nakatsuyama discloses:

A broadcast receiving method comprising:

a receiving step of receiving a broadcast in which additional information that is

made to correspond to an object appearing in broadcast contents broadcasted from a

broadcasting station and that contains keyword information specifying said object and a

language model are broadcasted simultaneously with said broadcast contents *("...providing a*

voice-recognition-based Internet shopping interface between an end user's Internet-enabled

device 10 and a plurality of online shopping sites 12, 14 and 16 via Internet connections 18...

communication via voice and the display of images (e.g., character images and scene images).

Internet-enabled device 10 can be, for example, an Internet-enabled personal digital assistant

(PDA), an Internet-enabled personal computer (PC), an Internet-enabled cellular phone, or an

interactive television...", "...predetermined agent interface functions can include, for example,

(i) voice recognition based on a Natural Language Interface (NLI) to convert the end user's

voice-format answers to text format answers; and (ii) a text-to-voice engine function for

converting text-format shopping site questions stored in the agent interface to voice-format

shopping site questions for presentation to the end user...", "...interactive dialog editor software

can, for example, present a retailer with a series of predetermined software-based forms 20 that

prompt a user to enter a series of text-format questions (e.g., Question 1), associated expected

answers (e.g., Ans11, Ansl2 etc.) and associated actions (e.g., Act11, Actl2 etc.) to be taken. For

ease of use, text-format product questions can be limited to a certain number (e.g., 10). As shown

in FIG. 3, the interactive dialog software can also prompt a retailer to supply pictures (e.g.,

Picture1, Picture2, etc.) of target products (e.g., Product1, Product2, etc.)…", Paragraphs

[0023], [0024], [0029], [0030], [0032], [0033], [0039], [0040]).

Nakatsuyama does not explicitly disclose the limitations a correcting step of utilizing a

synonym dictionary in which a plurality of words are classified into word classes on the basis of

the synonymy between the words, and of thereby correcting a frequency of appearance of a

predetermined combination of said word classes in an expression form of said language model

and/or a frequency of appearance of a predetermined word with reference to said word class in

an expression form of said language model, on the basis of history information of speech

recognition result of already performed speech recognition; a speech recognition step of

performing speech recognition of a voice uttered by a viewing person, by using said corrected

language model; a specifying step of specifying said keyword information on the basis of the

speech recognition result; and a displaying step of displaying additional information

corresponding to said specified keyword information. Mitchell, however, discloses the

limitations a correcting step of utilizing a synonym dictionary in which a plurality of words are

classified into word classes on the basis of the synonymy between the words, and of thereby

correcting a frequency of appearance of a predetermined combination of said word classes in an

expression form of said language model and/or a frequency of appearance of a predetermined

word with reference to said word class in an expression form of said language model, on the

basis of history information of speech recognition result of already performed speech recognition

("...a user model 21 which can be updated to improve the accuracy of the recognition, a

language model 22 and a dictionary 23 to which a user can add new words. The user model 21

comprises an acoustic model and a contextual model. During operation of the speech

recognition engine application 11 the application utilizes the user model 21, the language model

22 and the dictionary 23 in the memory 20 and outputs speech recognition data 24 to the

memory 20...", "...Each of the words recognized is identified by an identifier tag which

identifies the position in the sequence of word. Also, the audio start point and audio end point of

the audio component in the associated audio data file is indicated to enable the retrieval and

playback of the audio component corresponding to the word. For each word, a list of alternative

words and their scores is given where n is the score, i.e. the likelihood that the word is correct,

and w is the word. The list of alternative words is ordered such that the most likely word appears

first. Alternatives, if any, are then listed in order with the word having the highest score first and

the word having the lowest score last...", "...This information is provided in the tag field. The

tag field will not only include the identified tag identifying the position of the audio component

for a word within a file, it will also include an identification of which file contains the audio

component...", "...the error correction of step S11 of FIG. 5. In step S50 the user selects a word

which is believed to be incorrectly recognized for correction. The selected word is then

highlighted on the display in step S51... the speech recognition interface application 12

determines the word location in the text... speech recognition run time created files... a user can

select an alternative word from the choice list, input a new word, default back to the original

word or cancel if the original word is correct or the word was selected for correction in

error…", Col. 5, line 64- col. 6, line 7, Col. 7, lines 4-15, 20-24, 31-35, Col. 9, line 42 -Col. 10,

line 3);

a speech recognition step of performing speech recognition of a voice uttered by a

viewing person, by using said corrected language model *("…a user model 21 which can be*

updated to improve the accuracy of the recognition, a language model 22 and a dictionary 23 to

which a user can add new words. The user model 21 comprises an acoustic model and a

contextual model. During operation of the speech recognition engine application 11 the

application utilizes the user model 21, the language model 22 and the dictionary 23 in the

memory 20 and outputs speech recognition data 24 to the memory 20…", Col. 5, line 64- col. 6,

line 7);

a specifying step of specifying said keyword information on the basis of the speech

recognition result *("…Each of the words recognized is identified by an identifier tag which*

identifies the position in the sequence of word. Also, the audio start point and audio end point of

the audio component in the associated audio data file is indicated to enable the retrieval and

playback of the audio component corresponding to the word. For each word, a list of alternative

words and their scores is given where n is the score, i.e. the likelihood that the word is correct,

and w is the word. The list of alternative words is ordered such that the most likely word appears

first. Alternatives, if any, are then listed in order with the word having the highest score first and

the word having the lowest score last…", Col. 7, lines 4-15) ; and

a displaying step of displaying additional information corresponding to said specified

keyword information *("...a choice list can be displayed which comprises the alternative words*

listed alphabetically for ease of use. Corrections can then be carried out either by selecting one

of the alternative characters or entering a new character...", "...The selected word is then

highlighted on the display in step S51... the speech recognition interface application...", Col. 3,

lines 24-32, Col. 9, line 42 -Col. 10, line 3).

 Nakatsuyama and Mitchell are analogous art because they are from a similar field of

endeavor in speech recognition-based interface processing systems. Thus, it would have been

obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings

of Nakatsuyama with the correcting, recognizing, display and selection means taught by

Mitchell in order to advantageously permitting "...audio recording of the dictation stored which

can be replayed to aid the correction of the recognized text...providing an interface between the

output of a speech recognition engine and application capable of processing the output...to link

the relationship between the output data and the audio data...in such a way as to remove, reorder,

delete, insert or format the data", *(Col. 1, lines 46-48, Col. 2, lines 11-19).*

 With respect to **Claims 14, 17, 20, 26, 37**, Nakatsuyama discloses:

A broadcast receiving method comprising:

a receiving step of receiving a broadcast in which additional information that is made to

correspond to an object appearing in broadcast contents broadcasted from a broadcasting

station and that contains keyword information specifying said object and information specifying

a language model are broadcasted simultaneously with said broadcast contents *("...providing a*

voice-recognition-based Internet shopping interface between an end user's Internet-enabled

device 10 and a plurality of online shopping sites 12, 14 and 16 via Internet connections 18…

communication via voice and the display of images (e.g., character images and scene images).

Internet-enabled device 10 can be, for example, an Internet-enabled personal digital assistant

(PDA), an Internet-enabled personal computer (PC), an Internet-enabled cellular phone, or an

interactive television…", "…predetermined agent interface functions can include, for example,

(i) voice recognition based on a Natural Language Interface (NLI) to convert the end user's

voice-format answers to text format answers; and (ii) a text-to-voice engine function for

converting text-format shopping site questions stored in the agent interface to voice-format

shopping site questions for presentation to the end user…", "…interactive dialog editor software

can, for example, present a retailer with a series of predetermined software-based forms 20 that

prompt a user to enter a series of text-format questions (e.g., Question 1), associated expected

answers (e.g., Ans11, Ans12 etc.) and associated actions (e.g., Act11, Act12 etc.) to be taken. For

ease of use, text-format product questions can be limited to a certain number (e.g., 10). As shown

in FIG. 3, the interactive dialog software can also prompt a retailer to supply pictures (e.g.,

Picture1, Picture2, etc.) of target products (e.g., Product1, Product2, etc.)…", Paragraphs

[0023], [0024], [0029], [0030], [0032], [0033], [0039], [0040]).

Nakatsuyama does not explicitly disclose the limitations a language model specifying step of

specifying said language model retained in advance, by using information specifying said

received language model; a correcting step of utilizing a synonym dictionary in which a plurality

of words are classified into word classes on the basis of the synonymy between the words, and of

thereby correcting a frequency of appearance of a predetermined combination of said word

classes in an expression form of said specified language model and/or a frequency of appearance

of a predetermined word with reference to said word class in an expression form of said specified

language model, on the basis of history information of speech recognition result of already

performed speech recognition; a speech recognition step of performing speech recognition of a

voice uttered by a viewing person, by using said corrected language model; a specifying step of

specifying said keyword information on the basis of the speech recognition result; and a

displaying step of displaying additional information corresponding to said specified keyword

information.   Mitchell, however, discloses the limitations a language model specifying step of

specifying said language model retained in advance, by using information specifying said

received language model *("...During operation of the speech recognition engine application 11*

the application utilizes the user model 21, the language model 22 and the dictionary 23 in the

memory 20 and outputs speech recognition data 24 to the memory 20...", Col. 5, line 64- col. 6,

line 7);

    a correcting step of utilizing a synonym dictionary in which a plurality of words are

classified into word classes on the basis of the synonymy between the words, and of thereby

correcting a frequency of appearance of a predetermined combination of said word classes in an

expression form of said specified language model and/or a frequency of appearance of a

predetermined word with reference to said word class in an expression form of said specified

language model, on the basis of history information of speech recognition result of already

performed speech recognition *("...a user model 21 which can be updated to improve the*

accuracy of the recognition, a language model 22 and a dictionary 23 to which a user can add

new words. The user model 21 comprises an acoustic model and a contextual model. During

operation of the speech recognition engine application 11 the application utilizes the user model

21, the language model 22 and the dictionary 23 in the memory 20 and outputs speech

recognition data 24 to the memory 20...", "...Each of the words recognized is identified by an

identifier tag which identifies the position in the sequence of word. Also, the audio start point

and audio end point of the audio component in the associated audio data file is indicated to

enable the retrieval and playback of the audio component corresponding to the word. For each

word, a list of alternative words and their scores is given where n is the score, i.e. the likelihood

that the word is correct, and w is the word. The list of alternative words is ordered such that the

most likely word appears first. Alternatives, if any, are then listed in order with the word having

the highest score first and the word having the lowest score last...", "...This information is

provided in the tag field. The tag field will not only include the identified tag identifying the

position of the audio component for a word within a file, it will also include an identification of

which file contains the audio component...", "...the error correction of step S11 of FIG. 5. In

step S50 the user selects a word which is believed to be incorrectly recognized for correction.

The selected word is then highlighted on the display in step S51... the speech recognition

interface application 12 determines the word location in the text... speech recognition run time

created files... a user can select an alternative word from the choice list, input a new word,

default back to the original word or cancel if the original word is correct or the word was

selected for correction in error...", Col. 5, line 64- col. 6, line 7, Col. 7, lines 4-15, 20-24, 31-35,

Col. 9, line 42 -Col. 10, line 3);

a speech recognition step of performing speech recognition of a voice uttered by a

viewing person, by using said corrected language model *("...a user model 21 which can be*

updated to improve the accuracy of the recognition, a language model 22 and a dictionary 23 to

which a user can add new words. The user model 21 comprises an acoustic model and a

contextual model. During operation of the speech recognition engine application 11 the

application utilizes the user model 21, the language model 22 and the dictionary 23 in the

memory 20 and outputs speech recognition data 24 to the memory 20…", Col. 5, line 64- col. 6,

line 7);

a specifying step of specifying said keyword information on the basis of the speech

recognition result *("…Each of the words recognized is identified by an identifier tag which*

identifies the position in the sequence of word. Also, the audio start point and audio end point of

the audio component in the associated audio data file is indicated to enable the retrieval and

playback of the audio component corresponding to the word. For each word, a list of alternative

words and their scores is given where n is the score, i.e. the likelihood that the word is correct,

and w is the word. The list of alternative words is ordered such that the most likely word appears

first. Alternatives, if any, are then listed in order with the word having the highest score first and

the word having the lowest score last…", Col. 7, lines 4-15); and

a displaying step of displaying additional information corresponding to said specified

keyword information *("…a choice list can be displayed which comprises the alternative words*

listed alphabetically for ease of use. Corrections can then be carried out either by selecting one

of the alternative characters or entering a new character…", "…The selected word is then

highlighted on the display in step S51… the speech recognition interface application…", Col. 3,

lines 24-32, Col. 9, line 42 -Col. 10, line 3).

Nakatsuyama and Mitchell are analogous art because they are from a similar field of

endeavor in speech recognition-based interface processing systems. Thus, it would have been

obvious to a person of ordinary skill in the art, at the time of invention, to modify the teachings

of Nakatsuyama with the correcting, recognizing, display and selection means taught by

Mitchell in order to advantageously permitting "…audio recording of the dictation stored which

can be replayed to aid the correction of the recognized text…providing an interface between the

output of a speech recognition engine and application capable of processing the output…to link

the relationship between the output data and the audio data...in such a way as to remove, reorder,

delete, insert or format the data", *(Col. 1, lines 46-48, Col. 2, lines 11-19).*

With respect to **Claim 27**, Mitchell further discloses:

wherein the information specifying said language model is an ID imparted to said language

model in advance *("…Each of the words recognized is identified by an identifier tag which*

identifies the position in the sequence of word. Also, the audio start point and audio end point of

the audio component in the associated audio data file is indicated to enable the retrieval and

playback of the audio component corresponding to the word. For each word, a list of alternative

words and their scores is given where n is the score, i.e. the likelihood that the word is correct,

and w is the word. The list of alternative words is ordered such that the most likely word appears

first. Alternatives, if any, are then listed in order with the word having the highest score first and

the word having the lowest score last…", "…This information is provided in the tag field. The

tag field will not only include the identified tag identifying the position of the audio component

for a word within a file, it will also include an identification of which file contains the audio

component…", Col. 7, lines 4-15, 20-24, 31-35).

With respect to **Claim 28**, Mitchell further discloses:

wherein: the information specifying said language model is keyword information for language

model specification *("...Each of the words recognized is identified by an identifier tag which*

identifies the position in the sequence of word. Also, the audio start point and audio end point of

the audio component in the associated audio data file is indicated to enable the retrieval and

playback of the audio component corresponding to the word. For each word, a list of alternative

words and their scores is given where n is the score, i.e. the likelihood that the word is correct,

and w is the word. The list of alternative words is ordered such that the most likely word appears

first. Alternatives, if any, are then listed in order with the word having the highest score first and

the word having the lowest score last...", Col. 7, lines 4-15);

the keyword information for language model specification is imparted also to said language

model retained in advance *("..., the audio start point and audio end point of the audio*

component in the associated audio data file is indicated to enable the retrieval and playback of

the audio component corresponding to the word. For each word, a list of alternative words and

their scores is given where n is the score, i.e. the likelihood that the word is correct, and w is the

word. The list of alternative words is ordered such that the most likely word appears first.

Alternatives, if any, are then listed in order with the word having the highest score first and the

word having the lowest score last...", Col. 7, lines 4-15, 20-24, 31-35); and

said language model specifying part specifies said language model depending on the

degree of agreement of those keywords for language model specification *("...the likelihood that*

the word is correct, and w is the word. The list of alternative words is ordered such that the most

likely word appears first…", Col. 7, lines 4-15, 20-24, 31-35).

Also, Nakatsuyama discloses "…interactive dialog editor software can, for example,

present a retailer with a series of predetermined software-based forms 20 that prompt a user to

enter a series of text-format questions (e.g., Question 1), associated expected answers (e.g.,

Ans11, Ansl2 etc.) and associated actions (e.g., Act11, Actl2 etc.) to be taken. For ease of use,

text-format product questions can be limited to a certain number (e.g., 10). As shown in FIG. 3,

the interactive dialog software can also prompt a retailer to supply pictures (e.g., Picture1,

Picture2, etc.) of target products (e.g., Product1, Product2, etc.)…", *(Paragraphs [0023], [0024],*

[0029], [0030], [0032], [0033], [0039], [0040]).

With respect to **Claim 29**, Mitchell further discloses:

wherein: if said correcting part corrects a frequency of appearance of a predetermined

word with reference to a predetermined word class in an expression form of said language

model, wherein said history information contains a word recognized in said already performed

speech recognition and said correcting part extracts a word contained in said word class

containing the word corresponding to said keyword information (*"…This information is*

provided in the tag field. The tag field will not only include the identified tag identifying the

position of the audio component for a word within a file, it will also include an identification of

which file contains the audio component…", "…the error correction of step S11 of FIG. 5. In

step S50 the user selects a word which is believed to be incorrectly recognized for correction.

The selected word is then highlighted on the display in step S51… the speech recognition

interface application 12 determines the word location in the text… speech recognition run time

created files… a user can select an alternative word from the choice list, input a new word,

default back to the original word or cancel if the original word is correct or the word was

selected for correction in error…", Col. 5, line 64- col. 6, line 7, Col. 7, lines 4-15, 20-24, 31-35,

Col. 9, line 42 -Col. 10, line 3);

with respect to a word contained in said history information among the extracted words,

a frequency of appearance of the word with reference to said word class in an expression form

of said language model is increased *("…Each of the words recognized is identified by an*

identifier tag which identifies the position in the sequence of word. Also, the audio start point

and audio end point of the audio component in the associated audio data file is indicated to

enable the retrieval and playback of the audio component corresponding to the word. For each

word, a list of alternative words and their scores is given where n is the score, i.e. the likelihood

that the word is correct, and w is the word. The list of alternative words is ordered such that the

most likely word appears first. Alternatives, if any, are then listed in order with the word having

the highest score first and the word having the lowest score last…", Col. 7, lines 4-15, 20-24,

31-35, Col. 9, lines 6-14); and

with respect to a word not contained in said history information among the extracted

words, a frequency of appearance of the word with reference to said word class in an expression

form of said language model is decreased *("…the speech recognition engine application 11*

outputs speech recognition data 24 and stores the data in run time files in a temporary directory

of the disk storage 15. Also, the audio data is stored in parallel as a run time file in the

temporary directory in step S32. The speech recognition interface application 12 detects whether

the most likely words output from the speech recognition engine application 11 are firm or

infirm, i.e. whether the speech recognition engine application 11 has finished recognizing that

word or not in step S33. If the speech recognition engine application 11 has not finished

recognizing that word, a word is still output as the most likely, but this could change, e.g. when

contextual information is taken into consideration…", Col. 9, lines 6-14, Col. 7, lines 4-15, 20-

24, 31-35).

       With respect to **Claim 30**, Mitchell further discloses:

wherein: if said correcting part corrects a frequency of appearance of a predetermined

combination of said word classes in an expression form of said language model, wherein said

history information contains a word recognized in said already performed speech recognition

and said correcting means extracts a word class containing a word corresponding to said

keyword information (*"Each of the words recognized is identified by an identifier tag which*

identifies the position in the sequence of word …This information is provided in the tag field. The

tag field will not only include the identified tag identifying the position of the audio component

for a word within a file, it will also include an identification of which file contains the audio

component…", "…the error correction of step S11 of FIG. 5. In step S50 the user selects a word

which is believed to be incorrectly recognized for correction. The selected word is then

highlighted on the display in step S51… the speech recognition interface application 12

determines the word location in the text… speech recognition run time created files… a user can

select an alternative word from the choice list, input a new word, default back to the original

word or cancel if the original word is correct or the word was selected for correction in

error…", Col. 5, line 64- col. 6, line 7, Col. 7, lines 4-15, 20-24, 31-35, Col. 9, line 42 -Col. 10,

line 3);

with respect to said extracted word class, a frequency of appearance of a predetermined

combination of said word classes in an expression form of said language model is increased

("…Each of the words recognized is identified by an identifier tag which identifies the position

in the sequence of word. Also, the audio start point and audio end point of the audio component

in the associated audio data file is indicated to enable the retrieval and playback of the audio

component corresponding to the word. For each word, a list of alternative words and their

scores is given where n is the score, i.e. the likelihood that the word is correct, and w is the

word. The list of alternative words is ordered such that the most likely word appears first.

Alternatives, if any, are then listed in order with the word having the highest score first and the

word having the lowest score last…", Col. 7, lines 4-15, 20-24, 31-35, Col. 9, lines 6-14);

and with respect to a word class not extracted, a frequency that the word class appears after

a predetermined sequence of said word classes in an expression form of said language model is

decreased *("…the speech recognition engine application 11 outputs speech recognition data 24*

and stores the data in run time files in a temporary directory of the disk storage 15. Also, the

audio data is stored in parallel as a run time file in the temporary directory in step S32. The

speech recognition interface application 12 detects whether the most likely words output from

the speech recognition engine application 11 are firm or infirm, i.e. whether the speech

recognition engine application 11 has finished recognizing that word or not in step S33. If the

speech recognition engine application 11 has not finished recognizing that word, a word is still

output as the most likely, but this could change, e.g. when contextual information is taken into

consideration…", Col. 9, lines 6-14, Col. 7, lines 4-15, 20-24, 31-35).

Also, Nakatsuyama discloses "…associated actions to be taken and target products form

a branching pathway (or network) that is employed by the computer system to determine the end

user's product requirements and a target product that meets those requirements. FIG. 4 depicts a

branching pathway relationship between text-format product questions (e.g., Q, Q1, Q2, Q21

etc.), associated expected answers (e.g., Ans1, Ans2, Ans3, etc.), associated actions (e.g., Act1,

Act2, etc.) to be taken and target products (Product A, Product B, Product C, Product D etc.) in

accordance with one exemplary embodiment of the present invention. The branching pathway

relationship between these questions, answers and actions enables the computer system to

determine a target product that meets specified requirements, in particular, the end user's product

requirements as communicated via voice-format answers to voice-format product questions",

(Paragraph [0033], [0039], [0040]).

With respect to **Claim 31**, Mitchell further discloses:

wherein: if said correcting means corrects a frequency of appearance of a predetermined

combination of said word classes in an expression form of said language model, wherein said

history information contains a word class containing a word recognized in said already

performed speech recognition and said correcting part extracts a word class corresponding to

said keyword information *("…a choice list is built and displayed on the display. The choice list*

comprises the list of alternative words displayed alphabetically. In step S80 a user can select an

alternative word from the choice list, input a new word, default back to the original word, or

cancel if the original word is thought to be correct…", "Each of the words recognized is

identified by an identifier tag which identifies the position in the sequence of word ...This

information is provided in the tag field. The tag field will not only include the identified tag

identifying the position of the audio component for a word within a file, it will also include an

identification of which file contains the audio component...", "...the error correction of step S11

of FIG. 5. In step S50 the user selects a word which is believed to be incorrectly recognized for

correction. The selected word is then highlighted on the display in step S51... the speech

recognition interface application 12 determines the word location in the text... speech

recognition run time created files... a user can select an alternative word from the choice list,

input a new word, default back to the original word or cancel if the original word is correct or

the word was selected for correction in error...", Col. 10, lines 41-56, Col. 5, line 64- col. 6, line

7, Col. 7, lines 4-15, 20-24, 31-35, Col. 9, line 42 -Col. 10, line 3);

with respect to said extracted word class, a frequency of appearance of a predetermined

combination of said word classes in an expression form of said language model is increased

("...Each of the words recognized is identified by an identifier tag which identifies the position

in the sequence of word. Also, the audio start point and audio end point of the audio component

in the associated audio data file is indicated to enable the retrieval and playback of the audio

component corresponding to the word. For each word, a list of alternative words and their

scores is given where n is the score, i.e. the likelihood that the word is correct, and w is the

word. The list of alternative words is ordered such that the most likely word appears first.

Alternatives, if any, are then listed in order with the word having the highest score first and the

word having the lowest score last...", Col. 7, lines 4-15, 20-24, 31-35, Col. 9, lines 6-14);

and with respect to a word class not extracted, a frequency of appearance of a predetermined

combination of said word classes in an expression form of said language model is reduced

("...the speech recognition engine application 11 outputs speech recognition data 24 and stores

the data in run time files in a temporary directory of the disk storage 15. Also, the audio data is

stored in parallel as a run time file in the temporary directory in step S32. The speech

recognition interface application 12 detects whether the most likely words output from the

speech recognition engine application 11 are firm or infirm, i.e. whether the speech recognition

engine application 11 has finished recognizing that word or not in step S33. If the speech

recognition engine application 11 has not finished recognizing that word, a word is still output

as the most likely, but this could change, e.g. when contextual information is taken into

consideration...", Col. 9, lines 6-14, Col. 7, lines 4-15, 20-24, 31-35).

     Also, Nakatsuyama discloses "...associated actions to be taken and target products form

a branching pathway (or network) that is employed by the computer system to determine the end

user's product requirements and a target product that meets those requirements. FIG. 4 depicts a

branching pathway relationship between text-format product questions (e.g., Q, Q1, Q2, Q21

etc.), associated expected answers (e.g., Ans1, Ans2, Ans3, etc.), associated actions (e.g., Act1,

Act2, etc.) to be taken and target products (Product A, Product B, Product C, Product D etc.) in

accordance with one exemplary embodiment of the present invention. The branching pathway

relationship between these questions, answers and actions enables the computer system to

determine a target product that meets specified requirements, in particular, the end user's product

requirements as communicated via voice-format answers to voice-format product questions",

(Paragraph [0033], [0039], [0040]).

With respect to **Claim 32**, Mitchell further discloses:

comprising transmitting part for transmitting an instruction corresponding to a predetermined

operation to a predetermined transmission destination when the predetermined operation is

performed on said displayed additional information *("...analyze and determine an end user's*

(i.e., customer's) shopping site requirements. Such analysis and determination are accomplished

by presenting the end user with series of voice-format shopping site questions and receiving the

end user's voice-format answers thereto. The voice-format shopping site questions are presented

and the voice-format answers received via the end user's Internet-enabled device 10 and Internet

connections 18...", Paragraphs [0026], [0028], [0023], [0024], [0029], [0030], [0032], [0033],

[0039], [0040]).

With respect to **Claim 33**, Nakatsuyama further discloses:

wherein said additional information is goods sales information and/or services sales information,

and wherein said instruction corresponding to a predetermined operation is a request for

brochure or purchase instruction information concerning said goods and/or said service

("...interactive dialog editor software can, for example, present a retailer with a series of

predetermined software-based forms 20 that prompt a user to enter a series of text-format

questions (e.g., Question 1), associated expected answers (e.g., Ans11, Ansl2 etc.) and

associated actions (e.g., Act11, Actl2 etc.) to be taken. For ease of use, text-format product

questions can be limited to a certain number (e.g., 10). As shown in FIG. 3, the interactive

dialog software can also prompt a retailer to supply pictures (e.g., Picture1, Picture2, etc.) of

target products (e.g., Product1, Product2, etc.)…", Paragraphs [0023], [0024], [0029], [0030],

[0032], [0033], [0039], [0040]).

With respect to **Claim 34**, Mitchell further discloses:

wherein said language model retained in advance has been acquired in advance through a

network *("…a user model 21 which can be updated to improve the accuracy of the recognition,*

a language model 22 and a dictionary 23 to which a user can add new words. The user model 21

comprises an acoustic model and a contextual model. During operation of the speech

recognition engine application 11 the application utilizes the user model 21, the language model

22 and the dictionary 23 in the memory 20 and outputs speech recognition data 24 to the

memory 20…", "…The author work stations 100a, 100b and 100c are connected via a network

101 under the control of a network server 102 to an editor work station 103. The network 101

can comprise any conventional computer network such as an ethernet or token ring…",Col. 5,

line 64- col. 6, line 7, Col. 7, lines 4-15, 20-24, 31-35, Col. 9, line 42 -Col. 10, line 3, Col. 12,

lines 46-59).

Conclusion

10.   The prior art made of record and not relied upon is considered pertinent to applicant's

disclosure. Please see PTO-892 Form.

11.   Any inquiry concerning this communication or earlier communications from the

examiner should be directed to Edgar Guerra-Erazo whose telephone number is (571) 270-3708.

The examiner can normally be reached on M-F 7:30a.m.-5:00p.m. EST. If attempts to reach the

examiner by telephone are unsuccessful, the examiner's supervisor, David Hudspeth can be

reached on (571) 272-7843. The fax phone number for the organization where this application or

proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent

Application Information Retrieval (PAIR) system. Status information for published applications

may be obtained from either Private PAIR or Public PAIR. Status information for unpublished

applications is available through Private PAIR only. For more information about the PAIR

system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private

PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you

would like assistance from a USPTO Customer Service Representative or access to the

automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Edgar Guerra-Erazo/
Examiner, Art Unit 2626

/David R Hudspeth/
Supervisory Patent Examiner, Art Unit 2626